CLAIMS

1   1.   In a cluster of computing nodes having shared access
2   to one or more volumes of data storage using a parallel
3   file system, a method for managing the data storage,
4   comprising:
5   initiating a data management   (DM) application in
6   the   cluster   using   a   data   management   application
7   programming   interface   (DMAPI)   of   the   parallel   file
8   system;
9   receiving a request submitted to the parallel file
10   system on one of the nodes to perform an operation on a
11   file in one of the volumes of data storage;
12   obtaining a data management access right from the
13   DMAPI responsive to the request; and
14   performing   the   operation   on   the   file   using   the
15   access right.

1   2.   A method according to 1, wherein initiating the data
2   management application comprises creating a session of
3   the   data   management   application   on   a   session   node
4   selected from among the nodes in the cluster, and wherein
5   obtaining   the   data   management   access   right   comprises
6   obtaining the right at the session node.

1   3.   A method according to claim 2, wherein initiating
2   the data management application comprises initiating a
3   data migration application, so as to free storage space
4   on at least one of the volumes of data storage, and
5   wherein receiving the request comprises generating an
6   event responsive to the request, and wherein obtaining
7   the right at the session node comprises associating a DM
8   token with the right at the session node for use in
9   invoking a DMAPI function to be applied to the file and

10   associating the token with the event, and wherein
11   performing the operation comprises migrating data at a
12   plurality of the nodes simultaneously by presenting the
13   token in connection with the DMAPI function.

1   4.   A method according to claim 2, wherein receiving the
2   request comprises receiving an invocation of a file
3   operation submitted to the parallel file system by a user
4   application on a source node, and sending a notification
5   of a DM event to the session node responsive to the
6   request, and wherein obtaining the right at the session
7   node comprises processing the event at the session node
8   subject to the access right.

1   5.   A method according to claim 1, wherein obtaining the
2   data management access right comprises acquiring a data
3   management lock on the file, so as to restrict other data
4   management and file operations on the file while the lock
5   is held.

1   6.   A method according to claim 5, wherein the operation
2   is a data management operation, and wherein acquiring the
3   data management lock comprises holding the lock over a
4   sequence of multiple kernel calls in the parallel file
5   system.

1   7.   A method according to claim 5, wherein the operation
2   is a file operation, and wherein acquiring the data
3   management lock comprises holding the lock for a single
4   kernel call in the parallel file system.

1   8.   A method according to claim 7, wherein the file
2   operation is one of a plurality of file operations to be
3   performed on the file, and wherein acquiring the data
4   management lock comprises allowing the plurality of file

5 operations to hold respective data management locks
6 simultaneously without mutual conflict.

1 9. A method according to claim 5, wherein acquiring the
2 data management lock comprises acquiring an exclusive
3 lock.

1 10. A method according to claim 5, wherein acquiring the
2 data management lock comprises acquiring a shared lock.

1 11. A method according to claim 5, wherein acquiring the
2 data management lock comprises selecting the lock from a
3 table of locks provided for both file operations and data
4 management operations.

1 12. A method according to claim 11, wherein performing
2 the operation comprises calling a DMAPI function to
3 perform a data management operation, and wherein
4 acquiring the data management lock comprises acquiring,
5 in a course of executing the DMAPI function, one of the
6 locks provided for the file operations for the duration
7 of the DMAPI function, so as to enable calling the DMAPI
8 function without presenting a DM token.

1 13. A method according to claim 5, wherein acquiring the
2 data management lock comprises providing the data
3 management lock within a hierarchy of locks supported by
4 the parallel file system.

1 14. Computing apparatus, comprising:
2 one or more volumes of data storage, arranged to
3 store data; and

4 a plurality of computing nodes, linked to access the
5 volumes of data storage using a parallel file system, and
6 arranged so as to enable a data management (DM)
7 application to be initiated using a data management

8 application programming interface (DMAPI) of the parallel
9 file system, such that when a request submitted to the
10 parallel file system is received on one of the nodes to
11 perform an operation on a file in one of the volumes of
12 data storage, a data management access right is obtained
13 from the DMAPI responsive to the request, and the
14 operation on the file is performed using the access
15 right.

1 15. Apparatus according to 14, wherein the nodes are
2 arranged to initiate the data management application by
3 creating a session of the data management application on
4 a session node selected from among the nodes in the
5 cluster, and wherein the data management access right is
6 obtained at the session node.

1 16. Apparatus according to claim 15, wherein the data
2 management application comprises a data migration
3 application, which frees storage space on at least one of
4 the volumes of data storage, and wherein an event is
5 generated responsive to the request, causing the session
6 node to associate a DM token with the right for use in
7 invoking a DMAPI function to be applied to the file and
8 to associate the token with the event, and wherein data
9 are migrated at the plurality of the nodes simultaneously
10 by presenting the token in connection with the DMAPI
11 function.

1 17. Apparatus according to claim 15, wherein the request
2 comprises an invocation of a file operation submitted to
3 the parallel file system by a user application on a
4 source node, and wherein the nodes are arranged so that a
5 notification of a DM event is sent to the session node
6 responsive to the request, and wherein the event is

7 processed at the session node subject to the access
8 right.

1 18. Apparatus according to claim 14, wherein the data
2 management access right is obtained by acquiring a data
3 management lock on the file, so as to restrict other data
4 management and file operations on the file while the lock
5 is held.

1 19. Apparatus according to claim 18, wherein the
2 operation is a data management operation, and wherein the
3 data management lock is held over a sequence of multiple
4 kernel calls in the parallel file system.

1 20. Apparatus according to claim 18, wherein the
2 operation is a file operation, and wherein the data
3 management lock is held for a single kernel call in the
4 parallel file system.

1 21. Apparatus according to claim 20, wherein the file
2 operation is one of a plurality of file operations to be
3 performed on the file, and wherein the plurality of file
4 operations are allowed to hold respective data management
5 locks simultaneously without mutual conflict.

1 22. Apparatus according to claim 18, wherein the data
2 management lock comprises an exclusive lock.

1 23. Apparatus according to claim 18, wherein the data
2 management lock comprises a shared lock.

1 24. Apparatus according to claim 18, wherein the data
2 management lock is selected from a table of locks
3 provided for both file operations and data management
4 operations.

1 25. Apparatus according to claim 24, wherein the
2 operation comprises a DMAPI function called to perform a
3 data management operation, and wherein the data
4 management lock comprises one of the locks provided for
5 the file operations, which is acquired, in a course of
6 executing the DMAPI function, for the duration of the
7 DMAPI function, so as to enable calling the DMAPI
8 function without presenting a DM token.

1 26. Apparatus according to claim 18, wherein the data
2 management lock is provided within a hierarchy of locks
3 supported by the parallel file system.

1 27. A computer software product providing a data
2 management application programming interface (DMAPI) for
3 use in a cluster of computing nodes having shared access
4 to one or more volumes of data storage using a parallel
5 file system, the product comprising a computer-readable
6 medium in which program instructions are stored, which
7 instructions, when read by the computing nodes, cause a
8 data management (DM) application to be initiated using
9 the DMAPI, such that when a request submitted to the
10 parallel file system is received on one of the nodes to
11 perform an operation on a file in one of the volumes of
12 data storage, a data management access right is obtained
13 from the DMAPI responsive to the request, and the
14 operation on the file is performed using the access
15 right.

1 28. A product according to claim 27, wherein the
2 instructions cause the data management application to be
3 initiated by creating a session of the data management
4 application on a session node selected from among the

39879S3

5 nodes in the cluster, and wherein the data management
6 access right is obtained at the session node.

1 29. A product according to claim 28, wherein the data
2 management application comprises a data migration
3 application, which frees storage space on at least one of
4 the volumes of data storage, and wherein the instructions
5 cause an event to be generated responsive to the request,
6 causing the session node to associate a DM token with the
7 right for use in invoking a DMAPI function to be applied
8 to the file and to associate the token with the event,
9 and wherein data are migrated at the plurality of the
10 nodes simultaneously by presenting the token in
11 connection with the DMAPI function.

1 30. A product according to claim 28, wherein the request
2 comprises an invocation of a file operation submitted to
3 the parallel file system by a user application on a
4 source node, and wherein the instructions cause a
5 notification of a DM event to be sent to the session node
6 responsive to the request and cause the event to be
7 processed at the session node subject to the access
8 right.

1 31. A product according to claim 27, wherein the data
2 management access right is obtained by acquiring a data
3 management lock on the file, so as to restrict other data
4 management and file operations on the file while the lock
5 is held.

1 32. A product according to claim 31, wherein the
2 operation is a data management operation, and wherein the
3 data management lock is held over a sequence of multiple
4 kernel calls in the parallel file system.

IL9-2000-0042US1   46

1 33. A product according to claim 31, wherein the
2 operation is a file operation, and wherein the data
3 management lock is held for a single kernel call in the
4 parallel file system.

1 34. A product according to claim 33, wherein the file
2 operation is one of a plurality of file operations to be
3 performed on the file, and wherein the plurality of file
4 operations are allowed to hold respective data management
5 locks simultaneously without mutual conflict.

1 35. A product according to claim 31, wherein the data
2 management lock comprises an exclusive lock.

1 36. A product according to claim 31, wherein the data
2 management lock comprises a shared lock.

1 37. A product according to claim 31, wherein the data
2 management lock is selected from a table of locks
3 provided for both file operations and data management
4 operations.

1 38. A product according to claim 37, wherein the
2 operation comprises a DMAPI function called to perform a
3 data management operation, and wherein the data
4 management lock comprises one of the locks provided for
5 the file operations, which is acquired, in a course of
6 executing the DMAPI function, for the duration of the
7 DMAPI function, so as to enable calling the DMAPI
8 function without presenting a DM token.

1 39. Apparatus according to claim 31, wherein the data
2 management lock is provided within a hierarchy of locks
3 supported by the parallel file system.